Review of Matrix Theory

Notations:
A – capital bold denotes a matrix;
a – lower case bold is a vector;
a – lower case regular is a scalar;
a_{ij} - ij-element of A;
det(A) - determinant of A;
tr(A) - a trace of A;

Basics
Matrix A is defined by its elements a_{ij}:

\[
A = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\]  

(1)

Sometimes, elements of A are denoted as [A]_{ij}.

Sum of 2 matrices is defined element-wise:

\[
C = A + B \rightarrow c_{ij} = a_{ij} + b_{ij}
\]  

(2)

Product of matrices is defined as:

\[
C = A \cdot B \rightarrow c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}
\]  

(3)

Note that the product of A and B is defined only if the number of columns of A is the same as the number of rows of B, i.e. A and B are m×n and n×l matrices.

Determinant of a square n×n matrix det(A):

\[
det(A) = |A| = \sum_{k=1}^{n} a_{ik} (-1)^{i+k} M_{ik}
\]  

(4)

where M_{ik} is the minor of a_{ik}, i.e. the determinant of the submatrix of A, which is obtained by deleting i-th row and k-th column from A.

Example:

\[
A = \begin{bmatrix}
a_{11} & a_{12} \\
a_{21} & a_{22}
\end{bmatrix} \rightarrow det(A) = a_{11}a_{22} - a_{12}a_{21}
\]  

(5)

The transpose of A is defined as

\[
B = A^T \rightarrow b_{ij} = a_{ji}
\]  

(6)

i.e. row and column indexes are exchanged.

Complex conjugate operation is applied element-wise:

\[
B = A^* \rightarrow b_{ij} = a_{ij}^*
\]  

(7)

The Hermitian conjugate of A is

\[
B = A^+ = (A^T)^* \rightarrow b_{ij} = a_{ji}^*
\]  

(8)

Product of a matrix A and a scalar c is defined element-wise:

\[
B = c \cdot A \rightarrow b_{ij} = c \cdot a_{ij}
\]  

(9)
Some properties of transpose:
\[(AB)^T = B^T A^T, (AB)^+ = B^+ A^+\] (10)

Properties of det:
\[
\begin{align*}
\det(AB) &= \det(A)\det(B); \quad \det(c\cdot A) = c^n \det(A) \\
\det(A^T) &= \det(A); \quad \det(A^+) = (\det(A))^T;
\end{align*}
\] (11)
for square A and B. If det(A)=0, A is called singular.

Trace of a matrix is the sum of diagonal elements:
\[tr(A) = \sum_{i=1}^{n} a_{ii}\] (12)

Some properties of trace:
\[
\begin{align*}
tr(A+B) &= tr(A) + tr(B) \\
tr(AB) &= tr(BA) \\
tr(ABC) &= tr(CAB) = tr(BCA)
\end{align*}
\] (13)

Rank of a matrix is the number of linearly independent columns or rows. Some properties:
\[
\begin{align*}
\text{rank}(A+B) &\leq \text{rank}(A) + \text{rank}(B) \\
\text{rank}(AB) &\leq \min\left(\text{rank}(A), \text{rank}(B)\right)
\end{align*}
\] (14)

Vector a is a n×1 matrix:
\[a = \begin{bmatrix} a_1 & a_2 & \ldots & a_n \end{bmatrix}^T\] (15)

Sometimes it is called column vector.

Scalar product of two vectors a and b:
\[a^*b = \sum_{i=1}^{n} a_i^* b_i\] (16)

Frobenius or Euclidean norm (length) of a vector is:
\[|a| = \sqrt{a^* a} = \sqrt{\sum_{i=1}^{n} |a_i|^2}\] (17)

Similarly, Frobenius norm of a matrix:
\[
\|A\| = \left(\sum_{i=1}^{n} \sum_{j=1}^{m} |a_{ij}|^2\right)^{1/2} = \sqrt{tr(A^*A)}
\] (18)

Inverse of a n×n matrix:
\[B = A^{-1} \quad \text{if} \quad AB = BA = I\] (19)

I - identity matrix, \([I]_{ij} = \delta_{ij} = 1 \text{ if } i=j, 0 \text{ otherwise.}\)

If rank(A)<n, then det(A)=0 and the inverse does not exist → A is singular.
Some properties of the inverse:
\[
(AB)^{-1} = B^{-1}A^{-1} \quad \text{(20)}
\]
\[
\det(A^{-1}) = \frac{1}{\det(A)}
\]
\[
(A^T)^{-1} = (A^{-1})^T
\]
\[
(A^*)^{-1} = (A^{-1})^*
\]
if all the inverses exist.
The inverse of \( A \) can be calculated as
\[
A^{-1} = \frac{C^T}{\det(A)}, \quad c_{ij} = (-1)^{i+j} M_{ij} \quad \text{(21)}
\]
where \( M \) is the minor as before.

The matrix inversion lemma (MIL):
\[
(A + BCD)^{-1} = A^{-1} - A^{-1}B(DA^{-1}B + C^{-1})^{-1}DA^{-1} \quad \text{(22)}
\]
where \( A \) is \( n \times n \), \( B \) is \( n \times m \), \( C \) is \( m \times m \), \( D \) is \( m \times n \) and all the inverses are assumed to exist.

A special case of (22) is Woodbury’s identity:
\[
(A + xx^+)^{-1} = A^{-1} - \frac{A^{-1}xx^+A^{-1}}{1 + x^+A^{-1}x} \quad \text{(23)}
\]

Note: the product \( xx^+ \) is defined as
\[
B = xx^+ \rightarrow h_{ij} = x_i x_j^* \quad \text{(24)}
\]
i.e. element-wise.

Some special matrices

Symmetric matrix:
\[
A = A^T \rightarrow a_{ij} = a_{ji} \quad \text{(25)}
\]

Hermitian matrix:
\[
A = A^+ \rightarrow a_{ij} = a_{ji}^* \quad \text{(26)}
\]

Unitary matrix:
\[
UU^+ = I = U^+U \rightarrow U^{-1} = U^* \quad \text{(27)}
\]

Columns of a unitary matrix are orthogonal, \( u_i^*u_j = \delta_{ij} \).

Diagonal matrix \( A \):
\[
a_{ij} = 0 \quad \text{if } i \neq j \quad A = \text{diag} \left( a_{11}, a_{22}, \ldots, a_{nn} \right) \quad \text{(28)}
\]

Positive definite matrix:
\[
\text{if } \quad x^+Ax > 0 \quad \forall x \neq 0 \quad \text{(29)}
\]

Positive semi-definite matrix:
\[
\text{if } \quad x^+Ax \geq 0 \quad \forall x \neq 0 \quad \text{(30)}
\]

If a matrix is (semi)positive-definite, it is also Hermitian. The converse is not true in general.
Projection Matrices

Projection (indempotent) matrix:

\[ P^2 = P \]  

(31)

Further, we consider only Hermitian projection matrices,

\[ P^* = P. \]

Consider a linear vector space spanned by the columns of \( n \times m \) matrix \( V \),

\[ S = \text{span} \{V\} \]  

(32)

Assume columns of \( V \) are linearly-independent. Projection of \( x \) onto \( S \) is

\[ x_S = Px, \quad P = V(V^*V)^{-1}V^+ \]  

(33)

Projection of \( x \) onto \( S_\perp \) is

\[ x_{S\perp} = P_\perp x, \quad P_\perp = I - P \]  

(34)

where \( S_\perp \) is the space orthogonal to \( S \).

Eigenvalue Decomposition

Eigenvector of a \( n \times n \) matrix:

\[ Au = \lambda u \rightarrow (A - \lambda I)u = 0 \]  

(35)

where \( \lambda \) is an eigenvalue. Eigenvectors give “invariant” directions if \( A \) is considered as linear transformation.

Solution to

\[ |A - \lambda I| = 0 \]  

(36)

gives \( n \) eigenvalues \( \lambda \). There are \( n \) orthonormal eigenvectors.

Define:

\[ U = [u_1 \quad u_2 \ldots \quad u_n], \quad UU^+ = I \]

\[ A = \text{diag} [\lambda_1 \quad \lambda_2 \ldots \quad \lambda_n] \]

Then,

\[ A = U\Lambda U^+ = \sum_{i=1}^{n} \lambda_i u_i u_i^+ \]  

(37)

This is eigenvalue decomposition of \( A \).

Some properties

\[ \text{tr}(A) = \sum_{i=1}^{n} \lambda_i \]  

(38)

\[ \det(A) = \prod_{i=1}^{n} \lambda_i \]  

(39)

\[ A^{-1} = U\Lambda^{-1}U^+ = \sum_{i=1}^{n} \lambda_i^{-1} u_i u_i^+ \]  

(40)

Let \( \lambda(A) \) denotes the eigenvalues of \( A \). Then:

\[ \lambda(cA) = c\lambda(A), \quad c - \text{scalar} \]  

(41)

\[ \lambda(A^m) = \lambda^m(A), \quad m=1, 2, 3\ldots \]  

(42)

If \( A \) is Hermitian, \( A = A^* \), then \( \text{Im}\{\lambda(A)\} = 0 \). If \( A \) is positive definite, then \( \lambda_i(A) > 0 \).

\[ \lambda_i(I) = 1, \quad i = 1, 2\ldots n \]  

(43)

\[ \lambda_i(P) = 1, \quad i = 1\ldots k, \quad \lambda_i(P) = 0, \quad i = k + 1\ldots n \]  

(44)

where \( P \) is a projection matrix onto \( k \)-dimentional space.
\[
\lambda_{\max}(A) = \max_{x} [x^T A x], \quad |x| = 1
\]  
(45)

If \( A \) is singular, \( \text{rank}(A) = k < n \), "pseudoinverse" can be defined using non-zero eigenvalues only,

\[
A^{-1} = \sum_{i=1}^{K} \frac{1}{\lambda_i} u_i v_i^T
\]

**Singular Value Decomposition**

Arbitrary \( n \times m \) matrix \( A \) can be decomposed as

\[
A = U \Sigma V^+ = \sum_{i=1}^{l} \sigma_i u_i v_i^T
\]  
(46)

where \( U, V \) are unitary \( n \times n \) and \( m \times m \) matrices, and \( \Sigma \) is \( n \times m \) matrix,

\[
\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma_i = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_i)
\]  
(47)

where \( \sigma_i \geq 0 \) are singular values of \( A \), and \( u_i \) are the columns of \( U \) (the left singular vectors of \( A \)), \( v_i \) are the columns of \( V \) (the right singular vectors of \( A \)).

Note: singular values of \( A \) are non-negative square roots of the eigenvalues of \( A A^+ \). The right singular vectors of \( A \) are the eigenvectors of \( A^+ A \). Note from (46) that

\[
A V_K = \sigma_K u_K, \quad V_K^+ A = \sigma_K v_K^T
\]  
(49)

Pseudoinverse \( A^{-1} \) of a \( m \times n \) matrix \( A \) for \( m > n \) is defined from the following

\[
A^{-1} A = I_{n \times n}
\]  
(50)

where \( I_{n \times n} \) is \( n \times n \) identity matrix. Using the SVD of \( A \),

\[
A^{-1} = \sum_{i=1}^{n} \frac{1}{\sigma_i} u_i v_i^T = U \Sigma^{-1} V^+
\]  
(51)

where

\[
\Sigma^{-1} = \begin{bmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{bmatrix}.
\]

\( A^{-1} \) can be expressed as:

\[
A^{-1} = (A^+ A)^{-1} A^+
\]  
(52)

The above discussion assumes that \( A \) has the full column rank, i.e. linearly-independent columns. If \( n > m \) and \( A \) has full row rank, similar expressions hold true.

Pseudoinverse and projection matrix:

\[
P_A = AA^{-1}, \quad P_{\perp A} = I - AA^{-1}
\]  
(53)

**Properties of pseudoinverse**:

\[
A A^{-1} A = A, \quad A^{-1} A A^{-1} = A^{-1}
\]  
(54)

\[
(A^+)^{-1} = (A^{-1})^+
\]

\[
(A^+ A)^{-1} = A^{-1} (A^{-1})^+
\]  
(55)
\[(A^+A)^{-1} A^+ = A^{-1}\]

If \(B\) is invertible, then
\[\left(BA\right)^{-1} B = A^{-1}\] (56)

If \(a\) is a column vector, then
\[a^{-1} = \frac{a^+}{|a|^2}\] (57)

**Miscellaneous**

Let \(a(i)\) be \(i\)-th column of \(A\), and \(h_i^T\) be \(i\)-th row of \(B\), then
\[AB = \sum_{i=1}^{n} a(i)h_i^T\] (58)

**Null space** of a matrix \(A\) is a set of vectors \(x\) that satisfy
\[Ax = 0\] (59)

**Range** of a matrix \(A\) is a set of vectors \(y\) that satisfy
\[Ax = y\] (60)

for any \(x\). Note that
\[\text{rank}(A) = \text{dim}(y)\] (61)

where \(\text{dim}(y)\) is the dimensionality of the \(y\). Additionally,
\[\text{dim}(x) + \text{dim}(y) = n\] (62)

for \(n \times n\) matrix.

**Important property:**
\[\det(I_{n \times n} + AB) = \det(I_{m \times m} + BA)\] (63)

where \(A, B\) are \(n \times m\) and \(n \times m\) matrices, and \(I_{n \times n}\) is \(n \times n\) identity matrix.

**References**

**Brief reviews of matrices**


**Books**


1 strongly recommended to everybody interested in smart antennas, array processing, MIMO systems. Solid knowledge of matrix theory is essential for these fields.